

Citation for published version:

Caro, DH, Sandoval-Hernández, A & Lüdtke, O 2014, 'Cultural, social, and economic capital constructs in international assessments: an evaluation using exploratory structural equation modeling', *School Effectiveness and School Improvement*, vol. 25, no. 3, pp. 433-450. <https://doi.org/10.1080/09243453.2013.812568>

DOI:

[10.1080/09243453.2013.812568](https://doi.org/10.1080/09243453.2013.812568)

Publication date:

2014

Document Version

Peer reviewed version

[Link to publication](#)

Publisher Rights

Unspecified

This is an Author's Original Manuscript of an article published in a slightly modified form in the Journal *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice* (2013). [Copyright Taylor & Francis]. The published article is available online at <http://www.tandfonline.com/> [Article DOI: 10.1080/09243453.2013.812568]."
<http://dx.doi.org/10.1080/09243453.2013.812568>

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/254154691>

Cultural, social, and economic capital constructs in international assessments: an evaluation using exploratory structural equation modeling

ARTICLE *in* SCHOOL EFFECTIVENESS AND SCHOOL IMPROVEMENT · JULY 2014

Impact Factor: 0.8 · DOI: 10.1080/09243453.2013.812568

CITATIONS

7

READS

191

3 AUTHORS, INCLUDING:



[Daniel Caro](#)

University of Oxford

20 PUBLICATIONS 88 CITATIONS

[SEE PROFILE](#)



[Andres Sandoval-Hernandez](#)

University of Bath

23 PUBLICATIONS 30 CITATIONS

[SEE PROFILE](#)

**Cultural, Social and Economic Capital Constructs in International
Assessments: An Evaluation Using Exploratory Structural Equation Modeling**

"This is an Author's Original Manuscript of an article published in a slightly modified form in the Journal School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice (2013). [Copyright Taylor & Francis]. The published article is available online at <http://www.tandfonline.com/> [Article DOI: 10.1080/09243453.2013.812568]."
<http://dx.doi.org/10.1080/09243453.2013.812568>

Daniel Caro,
Centre for Educational Assessment,
Department of Education ,
University of Oxford ,
Oxford , UK

Andres Sandoval-Hernandez,
Research and Analysis Unit
International association for the Evaluation of Educational Achievement
Hamburg, Germany

Oliver Lüdtke,
Department of Psychology,
Humboldt University,
Berlin, Germany

Abstract

The article employs exploratory structural equation modeling (ESEM) to evaluate constructs of economic, cultural, and social capital in international large-scale assessment (LSA) data from the Progress in International Reading Literacy Study (PIRLS) 2006 and the Programme for International Student Assessment (PISA) 2009. ESEM integrates the theory-generating approach of exploratory factor analysis (EFA) and theory-testing approach of confirmatory factor analysis (CFA). It relaxes the zero-loading restriction in CFA, allowing items to load on different factors simultaneously, and it provides measurement invariance tests across countries not available in EFA. A main criticism of international LSA studies is the extended use of indicators poorly grounded in theory, like socioeconomic status, that prevent the study of mechanisms underlying associations with student outcomes. This article contributes to addressing this criticism by providing statistical criteria to evaluate the fit of well-defined sociological constructs with the empirical data.

Keywords: exploratory structural equation modeling; cultural capital; social capital; multi-group analysis; PIRLS; PISA.

International large-scale-assessments (LSA) studies of student achievement have been undertaken for some 60 years (Foshay, Thorndike, Hotyat, & Walker, 1962). Over time, these studies have had a significant impact on educational policy in aspects such as curriculum reforms, educational standards, and educational budget decisions, among other educational initiatives (Grek, 2009; Rutkowski & Rutkowski, 2010; Schwippert, 2007; Schwippert & Lenkeit, 2012). Their influential role in shaping educational policy is reflected by the continued emergence of international assessment studies in different regions, thematic areas, and for different target populations (Rochex, 2006). In addition to the well-established TIMSS, Trends in International Mathematics and Science Study; PIRLS, Progress in International Reading Literacy Study; and PISA, Programme for International Student Assessment; some newer regional or language-focused initiatives are being now carried out in Southern and Eastern Africa (SACMEQ), Latin America (UNESCO-LLECE), and French-speaking countries (PASEC). At risk of oversimplifying, the common goals of international LSAs are to measure and compare academic achievement outcomes of students through the administration of achievement tests, as well as to identify student, family, school, and community factors associated with student outcomes through a set of background questionnaires. The main assumption being that assessing and comparing education systems is the primary way to increase their effectiveness.

Despite the important role LSAs have had in advancing our understanding about factors associated to student outcomes and the influence this kind of studies have had on informing public policy around the world, significant criticisms regarding their theoretical, methodological, and policy commitments have fuelled a prolific debate about its boundaries and potentials. Both promoters (e.g., Teddlie & Reynolds, 2001) and detractors (e.g. Rochex, 2006; Slee, Weiner, & Tomlinson, 1998) of the use of international surveys concur that there is a lack of theory in the research program. These criticisms can be tracked down to the building blocks at the basis of the

architecture of the theories subscribed by LSA research. On the one hand they refer to the epistemological rationale behind the selection, operationalization, and explanation of the relationships among constructs (name deleted to maintain the integrity of the review process); and on the other to the rather simplistic ontology that underpins the fundamental assumptions of LSA research regarding the nature of schools, students, and teachers (Thrupp, 1999; Wrigley, 2004).

Authors like Coe and [Fitz-Gibbon \(1998\)](#) and [Lauder, Jamieson, and Wikeley \(1998\)](#) claim that in LSA research the decisions on what constructs are to be included in the surveys and its subsequent operationalization tends to relay on no more than common sense and statistical criteria, without considering the theories available in education and other disciplines. Following on these lines, Lauder and colleagues (op cit.) argue that the procedure used by LSA researchers to generate concepts like ‘socioeconomic status’ (SES) or ‘school climate’ is simply operationalize these concepts with a battery of items which when combined ‘ad up’ to proxy measures (name deleted to maintain the integrity of the review process). If they then turn to report significant coefficients in statistical models, they are regarded as important factors for improving education systems. A further problem of this kind of practice is that the battery of items used for operationalizing such concepts varies considerably in its content from one study to other, making it very difficult to accumulate the knowledge that would allow for theory construction (Chudgar, Luschei, Fagioli, & Lee, 2012).

A related criticism concerns the ‘fishing for correlations’ practice of testing for statistical relationships between particular constructs (e.g. SES or school climate) and student outcomes, without fully understanding why or how it is expected these two constructs to be related (Coe & Fitz-Gibbon, 1998). In addition, this kind of practices are normally coupled with causal interpretations of findings that follow a simple linear logic (e.g. more of this is associated to

more of that). Such practices oversimplify the complex nature of educational systems in which educational outcomes result from interactions and interrelated decisions of different educational actors (i.e., school staff, children, parents, etc.) ([Murnane & Willett, 2011](#); [Rochex, 2006](#)). From this perspective, we can claim to have gained great knowledge about the factors related to school effectiveness and yet, we know little about the mechanisms at work. As Murillo summarizes “... it can be said that we are beginning to know some things about what works in education, but the ignorance about *why* it works is still very important” (Murillo et al., 2007, p. 86).

Furthermore, the epistemological rationale of LSA research presupposes a particular social ontology that has also been questioned. Here, the criticisms concern what authors like [Willmott \(1999\)](#) termed as ‘a depthless ontology’, whereby what is ‘real’ is only what can be observed and measured; and educational actors are assumed to respond passively to incentives and sanctions of the system. These assumptions may not be adequate for the study of educational problems since, first, they ignore the agency of the actors (name deleted to maintain the integrity of the review process) and the context in which they are embedded ([Thrupp, 2001](#)); and second, because in essence, this perspective reduces scientific knowledge to what amounts to a description and/or prediction rather than a true explanation of social patterns ([Layder, 1990](#)).

Methodologically, construct development in LSA data has been limited by available techniques for theory generating and theory testing: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA makes it possible to examine the factor structure of construct items but not to test the invariance of constructs across countries. CFA allows for invariance tests, but often yields inflated interfactor correlations and model fit indices that fail to support constructs with well-defined EFA structures. This results, at least in part, from the overly restrictive assumption of CFA that construct items can load on a single factor only, that is, cross-loadings are fixed to zero (Marsh, Liem, Martin, Morin, & Nagengast, 2011). The inflated

interfactor correlations cast doubt about the differential validity of constructs and potentially introduce multicollinearity into models where the constructs are considered as predictors, distorting parameter estimates and obscuring the predicting ability of constructs (Marsh et al., 2011). Poor model fit drives researchers to modify CFA models in a process that seems more exploratory than confirmatory in nature.

Recently, the introduction of exploratory structural equation models (ESEM) has contributed to overcome some of the limitations of EFA and CFA (Marsh et al., 2009). ESEM integrates the statistical properties of EFA and CFA. Essentially, ESEM is an EFA model that provides standard errors, model fit indices, and measurement invariance tests. Whenever CFA yields lower goodness-of-fit indices and higher interfactor correlations than ESEM, ESEM is preferred. ESEM has already been applied in psychological research to evaluate multidimensional models of university teaching (Marsh et al., 2009); personality trait domains (Marsh et al., 2010; Rosellini & Brown, 2011) and how they vary across different survey conditions (e.g., telephone interviewing, face-to-face surveys, and self-administrated questionnaire) (Lang, John, Lüdtke, Schupp, & Wagner, 2011); motivation and engagement (Marsh et al., 2011); and post-traumatic stress disorders in the military after deployment (van Zuiden et al., 2011). To our knowledge, however, ESEM has not been yet applied to evaluate sociological constructs in international LSA data.

Purpose and Contribution

This paper applies ESEM to evaluate constructs of cultural, economic, and social capital in international LSA data from PIRLS 2006 and PISA 2009. The constructs are inspired in well-established theories by Bourdieu and [Passeron \(1977\)](#), [Bernstein \(1975\)](#), and [Coleman \(1988\)](#) and represent attempts to apply these complex theories in empirical studies. Factor structure of constructs is examined, assessing and comparing fit indices of CFA and ESEM models, and

testing for different levels of construct invariance across countries. The analyses represent an initial effort aimed at addressing some of the criticisms leveled at the theoretical underpinnings of international LSA studies. The paper aims to contribute to the literature in several respects.

First, it lays the foundation for the study of the mechanisms whereby families influence academic achievement. Previous research using SES indicators or single family background variables has been unable to explain the mechanisms at work for associations. Our theory informed analyses could provide researchers with a framework for studying ‘how’ family processes relate to student outcomes, and policy makers with sounder bases for designing interventions. Second, the paper helps detach the analysis and interpretation of statistical results from the simple linear logic attributed to traditional LSA research. Unlike the standard CFA model, the ESEM allows for construct items to load simultaneously on various constructs. In our empirical model, for example, the variable ‘number of books’ can load on both cultural capital and economic capital, reflecting more precisely the complexities of existing theoretical models (see Bourdieu, 1983). Fourth, the analysis of measurement invariance evaluates the validity of constructs across countries and offers researchers benchmarks for the possibilities of comparing results in analysis involving several countries.

Theoretical Framework

The literature is clear about the importance of cultural, social and economic capital constructs to student performance. How these constructs are defined and operationalized depends on the theoretical approach and the specific contexts. But the theories by Bourdieu and Passeron (1977), Bernstein (1975), and Coleman (1988) are usually regarded as a useful framework to delineate these concepts.

Bourdieu defines capital as “...accumulated labour (in its materialised form or its ‘incorporated’, embodied form) which, when appropriated on a private, i.e., exclusive basis by

agents or groups of agents, enables them to appropriate social energy in the form of reified or living labour” (Bourdieu, 1983, p. 241). The educational performance of students depends on the amount and composition of different forms of capital and on the extent to which these satisfy the symbolic requisites of the dominant culture legitimated by the education system (Bourdieu & Passeron, 1977). Bourdieu distinguishes between several types of capital. The concepts of cultural and economic capital are especially relevant for this work.

Cultural capital reflects the cultural long-lasting dispositions embedded in the human mind and body, as well as in cultural goods and educational credentials. It can appear in three states (Bourdieu, 1983): objectified, institutionalized and embodied. In its objectified state, cultural capital consists of cultural objects, such as pictures, books, didactic materials, instruments and machines such as a PC, or even the access to Internet. In its institutionalized state, cultural capital consists of educational credentials such as academic degrees held by parents. In its embodied form, cultural capital consists of permanent dispositions acquired and inherited through family socialization that are reflected in the access to cultural practices legitimated by the dominant culture (e.g., theatre, museum exhibitions, music concerts), the production of legitimate signals (e.g., a specific linguistic code), and the value attached to education (e.g., the amount of time spent reading for pleasure or academic aspirations).

Bernstein (1971) argued that the linguistic codes associated with parental occupational status reflected different forms of embodied cultural capital. He distinguished between restricted linguistic codes used by working class families and elaborated codes in middle-class families. Differences in the use of linguistic codes help explain achievement inequality. In particular, since schools are concerned with the introduction of new knowledge beyond existing shared meanings (Atherton, 2011), students who can handle an elaborated code (i.e., middle-class students) are more susceptible of performing better in the education system. For students who only have

restrictive codes at their disposal (i.e., students from disadvantaged contexts) it would be more difficult to interpret and produce the symbols legitimized in an education system thought for and by the dominant middle class (Blanco, 2007).

Economic capital is defined as the command students have over economic resources. This concept is commonly understood as exchange values, like income and assets that can be easily transformed into cash. Variables referring to the possession of consumer goods in the household are indicators of a family's economic capital, for example. Also, parental education and occupational status are usually regarded as resources that can be transformed into income. It is assumed that the greater the economic capital in a family the better the physical conditions supporting the cognitive development of its family members, and therefore the higher their educational attainment.

Theories by Bourdieu and Passeron (1977) and Bernstein (1971) have with no doubt extended our understanding of how cultural and economic resources are passed from parents on to children. But the deterministic nature of these so-called reproduction theories has prevented researchers from explaining some unusual patterns, such as academic resilience. And alternative explanations are offered when we consider social capital (e.g., Coleman, 1988) and rational action theory (e.g., Boudon, 1981), for example. We will use Coleman (1988) to frame the concept of social capital.

According to Coleman (1988) social capital is a resource that facilitates action. This concept takes rational action theory as a starting point, but rejects the extreme individualistic premises that often accompany it by taking account of the social structure. The link between reproduction theories and social capital is represented by those aspects of the social structure that can be used by actors as resources to achieve their interests, and by the decisions actors make regarding the use of these resources. Along these lines, social capital is represented by the

relationships among the family members that enhance the transmission of other structural resources like the parent's education, for example.

Method

Data

The data is sourced from PIRLS 2006 and PISA 2009 managed by the by International Association for the Evaluation of Educational Achievement (IEA) (Mullis, Martin, Kennedy, & Foy, 2007) and the Organisation for Economic Co-operation and Development (OECD) (OECD, 2010). PIRLS 2006 and PISA 2009 are the most recent international studies that assess reading achievement and gather information on student, family, and school aspects related to achievement. Target populations for PIRLS 2006 are 4th graders in 42 educational systems and for PISA 2009 are 15-year-olds in 65 educational systems. Family background data were collected through a parent questionnaire, which was optional in PISA 2009. The analyses using PISA 2009 are restricted to the 14 educational systems that adopted this option: Chile, Croatia, Denmark, Germany, Hong Kong, Hungary, Italy, Korea, Lithuania, Macao, New Zealand, Panama, Portugal, and Qatar.

For computational reasons, a random sample of 500 students per educational system was drawn from the population (OECD, 2012). Thus, estimates of ESEM models are based on 21,000 students (500x42) for PIRLS 2006 and 7,000 students (500x14) for PISA 2009. Overall, the original data of PIRLS 2006 and PISA 2009 included 36 and 66 items related to the economic, cultural, and social capital constructs, respectively. But final ESEM models include 14 selected items in PIRLS 2006 and 12 selected items for PISA 2009 (see Analytical Strategy).

Constructs and Selected Items

Economical Capital (F1)

Parental education. In PIRLS and PISA, a single variable reflecting the highest level of education of either parent. Response categories are 1-5 in PIRLS (1 “finished some primary or lower-secondary”, 2 “finished lower-secondary”, 3 “finished upper-secondary”, 4 “finished post-secondary education but not”, 5 “finished university or higher”) and 0-6 International Standard Classification of Education (ISCED) codes in PISA.

Parental occupational status. In PIRLS and PISA, the highest score of either parent on the International Socioeconomic Index of Occupational Status (ISEI; Ganzeboom, de Graaf, & Treiman, 1992). (name deleted to maintain the integrity of the review process) calculated a pseudo ISEI for PIRLS 2006.

Financial status. In PIRLS, parents' perception of the family financial situation. Responses categories are: 1 “not at all well-off”, 2 “not very well-off”, 3 “average”, 4 “somewhat well-off”, and 5 “very well-off”.

Children's own room. In PIRLS, a dichotomous indicator of own room possession (1 “yes”, 0 “no”).

Household overcrowding. In PISA, the proportion of household members to the number of rooms in the home.

Cultural Capital (F2)

Books at home. In PIRLS and PISA, the number of books at home excluding school related books. Response categories are 1-5 in PIRLS (1 “0-10 books”, 2 “11-25 books”, 3 “26-100 books”, 4 “101-200 books”, and 5 “more than 200 books”) and 1-6 in PISA (1 “0-10 books”, 2 “11-25 books”, 3 “26-100 books”, 4 “101-200 books”, 5 “201-500 books”, and 6 “more than 500 books”).

Children's books. In PIRLS, the number of children's books. Response categories are same as above (1-5).

Parental attitudes toward reading. In PIRLS, the index of parents' attitudes toward reading. Responses categories are: 1 "low", 2 "medium", and 3 "high".

Parents' read for enjoyment. In PIRLS, the frequency parents read for enjoyment. Responses categories are: 1 "never or almost never", 2 "once or twice a month", and 3 "once or twice a week", and 4 "every day or almost every day".

Parents' reading time. In PIRLS, the time parents spend reading at home. Responses categories are: 1 "less than one hour a week", 2 "1-5 hours a week", and 3 "6-10 hours a week", and 4 "more than 10 hours a week".

Parents reading motivation. In PISA, an index of motivational attributes of parents own reading engagement.

Cultural possessions. In PISA, a summary index reflecting possession of classical literature, books of poetry, and works of art.

Home educational resources. In PISA, a summary index reflecting possession of a desk to study, a quiet place to study, a computer for school work, educational software, a calculator, books to help with school work, and a dictionary.

Social Capital (F3)

Parents literacy support activities. In PIRLS and PISA, a summary index reflecting the frequency parents and children engaged in the following activities: read books, tell stories, sing songs, play with alphabet toys, talk about things they had done, talk about what they had read, play word games, write letters or words, and read aloud signs and labels.

Parent-child cultural communication. In PIRLS, parents reported the frequency they listened to the child read aloud and talk with the child about what he/she is reading on his/her own (responses categories are: 1 "never or almost never", 2 "once or twice a month", 3 "once or

twice a week”, and 4 “every day or almost every day”). Responses to these 2 items were averaged into a single index.

Parent-child communication and interaction. In PIRLS, the frequency parents talk to the child about things they have done. In PISA, a summary index of the frequency parents eat with the child around the table and the frequency parents spent time talking to the child.

Parent-child academic communication. In PIRLS and PISA, a summary index reflecting the frequency of parent-child communication about school related issues. Selected items in PIRLS are frequency of discussion about classroom reading work and the frequency parents helped the child with reading for school. Selected items in PISA are the frequency parents discussed with the child how well he/she was doing at school, the frequency parents talked with the child about what he/she was reading on his/her own, and the frequency parents helped the child with his/her homework.

Parent-child visit library/bookstore. In PIRLS and PISA, the frequency parents visited a library with the child (responses categories are: 1 “never or almost never”, 2 “once or twice a month”, 3 “once or twice a week”, and 4 “every day or almost every day”).

Parent-child social communication. In PISA, a summary index reflecting the frequency parents discussed films and political issues with the child.

Analytical Strategy

First, in unreported analyses we explored the structure of the complete social, cultural, and economic data (i.e., 36 items in PIRLS and 66 items in PISA) for dimensionality by means of traditional scree-plots and eigenvalues. The results yielded more than 3 factors and the interpretability of the rotated solution conveyed of item groups that captured sub-dimensions of theoretical constructs like embodied, objectified and institutionalized cultural capital. In order to achieve a 3 factor solution theoretically meaningful, item groups were summarized into single

indexes while trying to maintain approximately the same number of items per factor and dropping items that poorly measured factors. Ultimately, a well-fitting 3-factor solution was obtained with 14 items in PIRLS and 12 items in PISA.

Secondly, we compared CFA and ESEM models to demonstrate the advantages of ESEM. CFA imposed zero-loading restrictions on factor items, whereas ESEM allowed for items to load on the 3 factors simultaneously. Goodness-of-fit indices, interfactor correlations, and the interpretability of the rotated solution were compared. Goodness-of-fit was evaluated with the comparative fit index (CFI), Tucker–Lewis index (TLI), and root-mean-square error of approximation (RMSEA).¹ Models were estimated using the geomin rotation with an epsilon value of 0.5 and maximum likelihood estimation robust (MLR) in relation to non-normality and non-independence of observations. Geomin was selected to minimize factor complexity (i.e., smaller cross-factor loadings and larger inter-factor correlations) and make loadings more interpretable.²

Thirdly, multi-group analysis within the ESEM approach was conducted to evaluate factor and measurement invariance across countries in three steps following Meredith (1964) and Marsh, et al. (2010). The multi-group analysis is important because the same items and models are employed to reflect our sociological constructs in all the education systems analyzed, but the constructs might not remain constant across education systems. The analysis started by testing *configural invariance*, that is, the similarity of the overall pattern of parameters across countries with no restriction imposed on the parameters. The configural invariance model without invariance constraints serves as a baseline for comparing subsequent models that do impose equality constraints on the parameter estimates across countries. *Weak invariance* was tested next. The weak invariance test evaluates if factor loadings are invariant over countries. Then, the *strong measurement invariance* test evaluated if the indicator means (i.e., the intercepts of

responses to individual items) and factor loadings are invariant over countries. If strong measurement invariance is satisfied, changes in the latent factor means can be interpreted as changes in the latent constructs. *Strict measurement invariance* tests if factor loadings, item intercepts, and item uniquenesses (measurement error) are all invariant across groups. Strict measurement invariance is not tested because comparisons of latent constructs are purged of measurement error. Additionally, we performed country-wise weak invariance tests to indicate patterns of invariance across countries. All the analyses were performed with Mplus v. 6.1 (Muthén & Muthén, 2010).

Results

ESEM vs CFA

In applying ESEM we assumed, based on theory, that the latent factors would show significant cross-loadings and that compared to CFA, ESEM would fit the data better and yield lower interfactor correlations, that is, greater discriminant validity. The starting point of our analyses was thus to test whether this assumption holds. If it did not, then there would be no clear purpose for conducting ESEM.

Table 1 reports fit indices and Table 2 reports inter-factor correlations for the ESEM and CFA models applied to PIRLS and PISA. The indices in Table 1 for the CFA model did not support an acceptable fit to the data of PIRLS (e.g., TLI = 0.71; CFI = 0.77; RMSEA = 0.07) and PISA (e.g., TLI = 0.83; CFI = 0.87; RMSEA = 0.07). That is, the fit between the theoretical constructs and the empirical model is unsatisfactory according to the CFA specification. Furthermore, inter-factor correlations between economic capital and cultural capital for the CFA model (0.73 for PIRLS and 0.68 for PISA) exceeded acceptable levels under factor analysis with oblique rotation ($|0.50|$; Gorsuch, 1983). The size of inter-factor correlations was problematic for

distinguishing among theoretical constructs or supporting discriminant validity and casted doubts about the zero cross-loading restriction in CFA.

*** INSERT TABLE 1 ABOUT HERE ***

In contrast, ESEM results were acceptable in terms of both fit indices and inter-factor correlations. The ESEM solution (see Table 1) yielded acceptable fit indices for PIRLS (e.g., TLI = 0.93; CFI = 0.96; RMSEA = 0.03) and PISA (e.g., TLI = 0.93; CFI = 0.96; RMSEA = 0.05). In other words, the variables and the ESEM empirical model employed to operationalize the theoretical constructs provide a good fit to the data. Moreover, the inter-factor correlations were in all cases smaller than for the CFA solution. The highest inter-factor correlation of 0.41 for the economic capital and cultural capital factors in PIRLS was still within the acceptable range. The ESEM inter-factor correlations provided stronger support for discriminant validity than the CFA solution and likely indicated that the restriction of cross-loadings to be zero in CFA can distort the size of inter-factor correlations (Marsh et al., 2009). Overall, the results indicated that the ESEM model is more appropriate than CFA for evaluating constructs of economic, cultural, and social capital in the international assessments in terms of both model fit and discriminant validity.

*** INSERT TABLE 2 ABOUT HERE ***

Next, we examined factors loadings of the ESEM and CFA solution for the 14 items in PIRLS (see Table 3) and 12 items in PISA (see Table 4). The selected items loaded significantly and with the expected sign on their corresponding factors. The independent cluster CFA model restricted cross-loading correlations to zero, but the ESEM solution indicated that some items load significantly on more than one factor. For example, the item of parents literacy activities in PIRLS loaded positively and significantly on the social capital factor, as expected by theory, but also loaded significantly on the economic capital and cultural capital factors, though, with a

smaller loading size (see Table 3). Similarly, the parents reading motivation items in PISA loaded on its theoretical factor of cultural capital, but also on the economic capital and social capital factors (see Table 4). And, the parental occupational status item in PISA loaded significantly on the anticipated economic capital factor, but also on the cultural capital factor.

*** INSERT TABLES 3 AND 4 ABOUT HERE ***

Furthermore, the books at home item and children's books item in PIRLS (see Table 3) loaded greater on the cross-factor of economic capital (0.65 and 0.70, respectively) than on the postulated cultural capital factor (0.27 and 0.16, respectively). Cross-loading correlations for these items were higher than 0.3 and thus complicate the interpretation of the rotated solution ([Sass & Schmitt, 2010](#)). If driven solely by the data, the EFA component of ESEM would advice us to re-specify our CFA to include these items in the economic capital factor and not in the cultural capital factor. But PISA cross-loadings underlined a less complex structure: the cross-loading of the number of books item on the economic capital factor amounted to 0.23 and the loading on the theorized cultural capital factor amounted to 0.58 (see Table 4). And the reviewed theories offered solid bases for the postulated models. While the CFA model was not re-specified, the ESEM approach did reveal that the number of books item is indicative of both economic capital and cultural capital.

Multi-Group Analysis

We then evaluated factor and measurement invariance across educational systems using the ESEM model. The results are reported in Table 5. The baseline model without invariance constraints supported the configural invariance of our PISA model (TLI = 0.96; CFI = 0.98; RMSEA = 0.04) but could not be estimated for PIRLS. The PISA results indicated that the same model is able to fit the data for each group (educational system) when no constraints across groups are imposed. The model for PIRLS did not converge most likely due to the large number

of freely estimated parameters. Note that the number of groups in PIRLS, that is, 42 educational systems, is much larger than the 14 groups in PISA and that ESEM models seem to become unmanageable for larger groups (Marsh et al., 2009). Yet, it was possible to conduct subsequent multi-group analyses for both PIRLS and PISA.

When we started to impose constraints across education systems, model fit indices declined substantially (see Table 5). For example, the weak invariance model, the one constraining loadings to be invariant across education systems, yielded fit indices far from being acceptable for PIRLS (TLI = 0.55; CFI = 0.58; RMSEA = 0.13) and PISA (TLI = 0.45; CFI = 0.42; RMSEA = 0.14). In other words, there was no support for equivalence of loadings across education systems. This is the most basic level of invariance, which is a precondition for examining more stringent levels of invariance (e.g., strong and strict invariance). It is expected that if a construct does not satisfy weak invariance conditions it also will not satisfy strong and strict invariance conditions that impose even more restrictions on factor parameters across education systems in addition to restricting the factor loadings to be equal across education systems. Accordingly, the fit indices for the strong and strict invariance tests were also far from being acceptable (see Table 5).

But although factor and measurement invariance was not supported for the total number of education systems when included in combination, there might still be pairs of education systems where invariance conditions are met. As a form of exercise, we performed country-wise tests of weak invariance to identify patterns of invariance across education systems.

*** INSERT TABLE 5 ABOUT HERE ***

Table 6 reports selected fit indices (i.e., CFI, TLI, and RMSEA) of weak invariance tests for pairs of educational systems in PISA and summary statistics (i.e., median, minimum, and maximum) of fit indices. The matrix records 91 invariance tests corresponding to all possible

pair combinations of 14 educational systems. The matrix diagonal records ESEM model fit indices for individual educational systems. Fit indices that indicated reasonable support for weak invariance among two educational systems were highlighted. The 42 education systems in PIRLS would yield 861 invariance tests. For ease of interpretation, only PISA results are reported.

Mainly, the comparisons indicated that loadings were not equivalent for most pairs of education systems. The median values of CFI, TLI, and RMSEA amounted to 0.71, 0.63, and 0.11, respectively (see Table 6). Fit indices were not as low as the ones obtained when evaluating weak invariance for the 14 systems simultaneously but still were not acceptable. Further, some comparisons produced very poor weak invariance fit indices. For example, Macao and New Zealand (TLI = 0.32; CFI = 0.46; RMSEA = 0.15) produced the minimum values of TLI/CFI and Chile and Hong Kong (TLI = 0.39; CFI = 0.52; RMSEA = 0.16) the maximum value of RMSEA. Korea and Denmark (TLI = 0.35; CFI = 0.49; RMSEA = 0.14), New Zealand and Hong Kong (TLI = 0.41; CFI = 0.53; RMSEA = 0.14), and Panama and Korea (TLI = 0.41; CFI = 0.53; RMSEA = 0.15) also reported very poor values. In contrast, Lithuania and Croatia (TLI = 0.90; CFI = 0.92; RMSEA = 0.06) reported greatest support for weak invariance. Italy and Germany (TLI = 0.89; CFI = 0.91; RMSEA = 0.06), Italy and Hungary (TLI = 0.89; CFI = 0.91; RMSEA = 0.07), and Portugal and Italy (TLI = 0.89; CFI = 0.91; RMSEA = 0.06) also indicated reasonable support for weak invariance.

*** INSERT TABLE 6 ABOUT HERE ***

Discussion

International LSA studies using SES indicators have made significant contributions to the study of educational inequality, but are also limited by the formative nature of SES models (name deleted to maintain the integrity of the review process). SES adheres to a formative

measurement model because it depends mainly on its operationalization procedure and lacks a solid theoretical meaning (Bollen & Lennox, 1991). And, as such, its use for explaining the mechanisms for family influences on student performance is unclear. Instead, reflective indicators grounded in well-established theories, like social capital, cultural capital, and economic capital can illuminate the reasons for student performance, but are less frequently used in empirical studies. Methodologically, Principal Component Analysis wherein causality flows from the items into the construct adapts well to the SES formative measurement, whereas EFA and CFA models wherein causality flows from the construct into the items are appropriate for reflective models (name deleted to maintain the integrity of the review process). Recently, ESEM integrated some of the properties of EFA and CFA (Marsh et al., 2009).

The presented analyses evaluated constructs of cultural, social, and economic capital in the data sets of PIRLS and PISA with ESEM. The constructs were defined and operationalized using influential and complex theories by Bourdieu and Passeron (1977), Bernstein (1975), and Coleman (1988) as a conceptual guide. The results suggested that these constructs are reflected in the data sets and that ESEM models are superior to CFA models to evaluate them, as ESEM produced higher model fit indices and reasonable inter-factor correlations. The study of the association of these constructs with student performance could increase our understanding of the mechanisms whereby family background affects schooling outcomes. We therefore believe that the adoption the ESEM approach in educational models deserves the attention of further research.

The flexibility of ESEM in terms of model specification is consistent with the constructs concerned. For example, it was expected that cultural possessions would relate to both economic and cultural capital and that parental literacy activities with children would relate to both cultural and social capital. Accordingly, ESEM results indicated that the number of books loads

significantly on cultural and economic capital. In fact, in PIRLS, books exerted a greater load on economic capital than on its cultural capital corresponding construct, thereby suggesting that its extensive use as economic capital indicator might be justified. But PISA results indicated that the number of books loading is greater for cultural capital than for economic capital. This is in agreement with the cultural capital theory. And probably as access to Internet and online books increases, books will tend to indicate more cultural capital and less economic capital.

Also, as expected from the theory, parents literacy activities proved a good indicator of cultural and social capital. Similarly, parental motivation items loaded simultaneously on these two forms of capital. Bernstein (1971) suggested that parental occupational status reflected a form of embodied cultural capital. Therefore we expected occupational status not only to reflect economic capital but also cultural capital. Indeed, the results showed that the parental occupational status indicated mainly economic capital but also significantly cultural capital both in PISA and PIRLS. It could be then, as postulated by Bernstein (1971), that families with low prestige occupations (e.g., working class) develop restricted linguistic codes because they tend to concentrate in culturally homogeneous areas and are employed in jobs that demand low verbal elaboration, whereas families with higher prestige occupations (e.g., middle class) develop elaborated linguistic codes because they are geographically, socially and culturally more mobile and are employed in jobs that foster oral and written communication skills to a greater extent. Since schools recognize and value elaborated linguistic codes, differences in embodied cultural capital related to parental occupational status will translate in differences in student performance.

We were unable to validate measurement invariance of these constructs with the multi-group analysis. The model results did not support the most basic weak invariance assumption. In other words, we cannot claim that the constructs are reflected equally across nations. And comparisons of construct scores and effects across nations could be misleading. This might

reflect statistical complexities for managing invariance tests among a large number of groups. But even the country-wise analysis yielded low fit indices for weak invariance tests. Only in few pairs of countries the sociological constructs were somewhat comparable. Socioeconomic characteristics of countries might explain construct invariance results across countries. (name deleted to maintain the integrity of the review process) have shown, for example, that home possessions tend to load differently on SES across countries, more strongly in poorer countries and weakly in wealthier societies. And May (2006) has proposed a technique to enhance construct invariance by using the information of national specific items of home possessions. Further research should study more thoroughly the reasons for construct invariance across countries and propose methods to lessen construct variance.

Further research should also evaluate the presence of competing theories in the data sets. We have applied ESEM to evaluate constructs inspired by reproduction and social capital theories in international LSA, but we could have also evaluated constructs based on rational action theory (e.g., Boudon, 1981) and whether the fit for these models is better than the one for the reproduction theory. More generally, it could be possible to add an ontological dimension to international LSA studies, through which best explanations for observed phenomena at a given time and in specific contexts are sought. The knowledge gained from such an approach could result in sounder theory grounded comparisons among education systems.

Notes

1. Heuristically, values of TLI and CFI greater than 0.90 and 0.95, respectively, reflect acceptable and excellent fit to the data while for the RMSEA, values less than 0.05 and 0.08 reflect a close fit and a reasonable fit to the data (Marsh, Hau, & Wen, 2004).

2. Higher inter-factor correlations introduce problems for supporting discriminant validity and it is recommended not to employ factor analysis for inter-factor correlations larger than $|0.5|$ (Gorsuch, 1983). Cross-loading magnitudes greater than $|0.3|$ are indicative of complex factor structures and complicate the interpretation of the rotated solution (Sass & Schmitt, 2010).

References

- Atherton, J. S. (2011). Language Codes. *Doceo*. Retrieved June 5, 2012, from http://www.doceo.co.uk/background/language_codes.htm.
- Bernstein, B. B. (1971). *Class, codes and control. Theoretical studies towards a sociology of language*. (Vol. 1). London [etc.]: Routledge & Kegan Paul.
- Bernstein, B. B. (1975). *Class, codes and control. Towards a theory of educational transmissions* (Vol. 3). London: Routledge.
- Blanco, E. (2007). *Eficacia Escolar en Mexico. Factores escolares asociados a los aprendizajes en la educacion primaria* (Doctorate). Facultad Latinoamericana de Ciencias Sociales, Mexico, D.F.
- Bollen, K., & Lennox, R. (1991). Conventional Wisdom on Measurement: A Structural Equation Perspective. *Psychological Bulletin*, 110(2), 305–314. doi:10.1037/0033-2909.110.2.305.
- Boudon, R. (1981). *The logic of social action : an introduction to sociological analysis*. London; Boston: Routledge & Kegan Paul.
- Bourdieu, P. (1983). The Forms of Capital. In J. G. Richards (Ed.), *Readings in Economic Sociology* (pp. 241–258). New York: Greenwood Press.
- Bourdieu, P., & Passeron, J.-C. (1977). *Reproduction in education, society and culture*. Theory, Culture & Society. London: Sage Publications, Inc.
- Chudgar, A., Luschei, T. F., Fagioli, L. P., & Lee, C. (2012, April). *Socio-economic status (SES) measures using the Trends in International Mathematics and Science Study data*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, Canada.

- Coe, R., & Fitz-Gibbon, C. T. (1998). School Effectiveness Research: Criticisms and Recommendations. *Oxford Review of Education*, 24(4), 421–438.
- Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, 94, S95–S120. doi: 10.1086/228943.
- Foshay, A., Thorndike, R., Hotyat, P., & Walker, D. (1962). *Educational Achievement of Thirteen-Year-Olds in Twelve Countries*. Hamburg: UNESCO Institute for Education.
- Ganzeboom, H.B.G., de Graaf, P.M., & Treiman, D.J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21, 1-56. doi: 10.1016/0049-089X(92)90017-B.
- Gorsuch, R. L. (1983). *Factor analysis*. London: Lawrence Erlbaum.
- Grek, S. (2009). Governing by the numbers: the PISA ‘effect’ in Europe. *Journal of Education Policy*, 24(1), 23–27. doi: 10.1080/02680930802412669.
- Lang, F. R., John, D., Lüdtke, O., Schupp, J., & Wagner, G. G. (2011). Short assessment of the Big Five: robust across survey methods except telephone interviewing. *Behavior Research Methods*, 43(2), 548–567. doi:10.3758/s13428-011-0066-z.
- Lauder, H., Jamieson, I., & Wikeley, F. (1998). Models of Effective Schools: Limits and Capabilities. In R. Slee, G. Weiner, & S. Tomlinson (Eds.), *School Effectiveness for whom?: Challenges to the School Effectiveness and School Improvement Movements*. London: Falmer Press.
- Layder, D. (1990). *The Realist Image in Social Science*. London: Macmillan.
- Marsh, H. W., Hau, K. T., & Wen, Z. L. (2004) In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralising Hu & Bentler (1999) findings. *Structural Equation Modeling*, 11(3), 320-341. doi: 10.1207/s15328007sem1103_2.

- Marsh, H. W., Liem, G. A. D., Martin, A. J., Morin, A. J. S., & Nagengast, B. (2011). Methodological Measurement Fruitfulness of Exploratory Structural Equation Modeling (ESEM): New Approaches to Key Substantive Issues in Motivation and Engagement. *Journal of Psychoeducational Assessment*, 29(4), 322–346. doi:10.1177/0734282911406657.
- Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A New Look at the Big Five Factor Structure Through Exploratory Structural Equation Modeling. *Psychological Assessment*, 22(3), 471–491. doi:10.1037/a0019227.
- Marsh, H., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A., & Trautwein, U. (2009). Exploratory Structural Equation Modeling, Integrating CFA and EFA: Application to Students' Evaluations of University Teaching. *Structural Equation Modeling A Multidisciplinary Journal*, 16(3), 439–476. doi:10.1080/10705510903008220.
- May, H. (2006). A multilevel Bayesian item response theory method for scaling. *Journal of Educational and Behavioral Statistics*, 31(1), 63–79. doi: 10.3102/10769986031001063.
- Meredith, W. (1964). Notes on factorial invariance. *Psychometrika*, 2(2), 177–185. doi: 10.1007/BF02289699.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 International Report : IEA's Progress in International Reading Literacy Study in primary schools in 40 countries*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Murillo, F. J., Castañeda, E., Cueto, S., Donoso, J. M., Fabara, E., Hernandez, M. L., Herrera, M., et al. (2007). *Investigación iberoamericana sobre eficacia escolar [Ibero American school effectiveness study]*. Bogotá: Convenio Andrés Bello.

- Murnane, R. J., & Willett, J. B. (2011). *Methods matter: Improving causal inference in educational and social research*. New York: Oxford University Press.
- Muthén, L., & Muthén, B. (2010). *MPLUS [Computer Software]*. Los Angeles, CA: Muthén & Muthén.
- OECD. (2010). *PISA 2009 results. What Students Know and Can Do: Student Performance in Reading, Mathematics and Science* (Vol. I). Paris: OECD. Retrieved June 4, 2012 from <http://www.pisa.oecd.org/dataoecd/10/61/48852548.pdf>
- OECD (2012). *PISA 2009 Technical Report*. Paris: OECD Publishing. Retrieved June 4, 2012 from <http://dx.doi.org/10.1787/9789264167872-en>
- Rochex, J.-Y. (2006). Social, methodological, and theoretical issues regarding assessment: Lessons from a secondary analysis of PISA 2000 literacy tests. *Review of research in education*, 30(1), 163–212. doi: 10.3102/0091732X030001163 .
- Rosellini, A. J., & Brown, T. A. (2011). The NEO Five-Factor Inventory: Latent Structure and Relationships With Dimensions of Anxiety and Depressive Disorders in a Large Clinical Sample. *Assessment*, 18(1), 27 –38. doi:10.1177/1073191110382848.
- Rutkowski, L., & Rutkowski, D. (2010). Getting it ‘better’: the importance of improving background questionnaires in international large-scale assessment. *Journal of Curriculum Studies*, 42(3), 411–430. doi: 10.1080/00220272.2010.487546.
- Sass, D. A., & Schmitt, T. A. (2010). A Comparative Investigation of Rotation Criteria within Exploratory Factor Analysis. *Multivariate Behavioral Research*, 45(1), 73–103. doi:10.1080/00273170903504810.
- Schmitt, T. A. (2011). Current Methodological Considerations in Exploratory and Confirmatory Factor Analysis. *Journal of Psychoeducational Assessment*, 29(4), 304–321. doi:10.1177/0734282911406653

- Schwippert, K. (2007). *Progress in Reading Literacy: The Impact of PIRLS 2001 in 13 Countries*. Münster: Waxmann.
- Schwippert, K. & Lenkeit, J. (Eds.) (2012). *Progress in reading literacy in national and international context. The impact of PIRLS 2006 in 12 countries*. Münster, Germany: Waxmann.
- Slee, R., Weiner, G., & Tomlinson, S. (1998). *School Effectiveness for whom? Challenges to the school effectiveness and school improvement movements*. London: Falmer Press.
- Teddlie, C., & Reynolds, D. (2001). Countering the Critics: Responses to Recent Criticisms of School Effectiveness Research. *School Effectiveness and School Improvement*, 12(1), 41–82.
- Thrupp, M. (1999). *Schools Making a Difference: Let's Be Realistic!* London: Open University Press.
- Thrupp, M. (2001). Recent School Effectiveness Counter-critiques: problems and possibilities. *British Educational Research Journal*, 27(4), 443–457. doi: 0.1080/01411920120071452.
- van Zuiden, M., Geuze, E., Willemen, H. L. D. M., Vermetten, E., Maas, M., Heijnen, C. J., & Kavelaars, A. (2011). Pre-existing high glucocorticoid receptor number predicting development of posttraumatic stress symptoms after military deployment. *The American Journal of Psychiatry*, 168(1), 89–96. doi:10.1176/appi.ajp.2010.10050706.
- Willmott, R. (1999). Structure, Agency and the Sociology of Education: Rescuing Analytical Dualism. *British Journal of Sociology of Education*, 20(1), 5–21.
- Wrigley, T. (2004). 'School effectiveness': the problem of reductionism. *British Educational Research Journal*, 30(2), 227–244. doi: 10.1080/0141192042000195272.